(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 831 148 A1

(12) EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:
25.03.1998 Bulletin 1998/13

(21) Application number: **97905440.0**

(22) Date of filing: **05.03.1997**

(51) Int. Cl.[6]: **C12N 15/12**, C12N 15/63, C07K 14/435, C12N 1/21, C12P 21/02, C12Q 1/68, C07K 16/18, C12P 21/08, A61K 38/17, G01N 33/53 // C12N1:21

(86) International application number:
PCT/JP97/00665

(87) International publication number:
WO 97/32982 (12.09.1997 Gazette 1997/39)

(84) Designated Contracting States:
DE FR GB IT

(30) Priority: **07.03.1996 JP 49880/96**
12.12.1996 JP 331944/96

(71) Applicant: **Eisai Co., Ltd.**
Tokyo 112-88 (JP)

(72) Inventor: **TSUKITA, Sholchiro**
Nakagyo-ku, Kyoto-shi, Kyoto 604 (JP)

(74) Representative:
Hansen, Bernd, Dr. Dipl.-Chem. et al
Hoffmann Eitle,
Patent- und Rechtsanwälte,
Arabellastrasse 4
81925 München (DE)

(54) HUMAN ADHESION MOLECULE OCCLUDIN

(57) Whole structures of mammalian analogues of occludin, a constituent protein of the tight junction (TJ), are provided.

Genes for human, canine and mouse occludins were analyzed with the PCR technique on the basis of the coding sequence seen around the gene for neuronal apoptosis inhibitory protein. With antibodies prepared, the occludins have been confirmed to be constituent proteins of the TJ by immunofluorescent cell staining.

EP 0 831 148 A1

Printed by Xerox (UK) Business Services
2.15.12/3.4

Description

Technical Field

The present invention relates to an amino acid sequence of membrane protein occludin in a tight junction (hereinafter referred to as "the TJ") of a human, a dog and a mouse, and a DNA for encoding the amino acid sequence.

Background Art:

In multicellular animals, the information on cellular adhesion between adjacent cells is deeply concerned with the regulation and maintenance of vital phenomena such as cellular proliferation and differentiation, inflammation and cancer metastasis. Intercellular adhesion molecules which take part in the adhesion of the cells frequently assemble together on the surfaces of these cells to form a specifically differentiated membrane region for the adhesion. Particularly, intercellular adhesion molecules such as cadherins are known to be firmly bound to cytoskeleton in the cytoplasmic domain of epithelial cells. Such membrane regions are called intercellular adhesion apparatus and chiefly classified into the following four structures: gap junction (GJ), adherens junction (AJ), desmosome and tight junction (TJ).

These adhesion apparatus has first been identified under an electron microscope, and as a result of the investigation and research of constituent proteins, the importance of their physiological and pathological significance has become the focus of increased interest. The proteins that are called the so-called adhesion molecules specifically exist in these adhesion apparatus, and the adhesion molecules of the AJ are cadherins and various kinds of cadherins such as N-cadherin and P-cadherin have been identified so far [Takeichi, M. et al., "Science", Vol. 251, p. 1451-1455, (1991)]. As adhesion molecules of the desmosome, desmoglein and desmocollin are known, and according to recent studies, it has been elucidated that their structures are similar to those of the cadherins [Buxton, R. S. et al., "J Cell Biol.", Vol. 121, p. 481-484, (1993)]. The adhesion molecules of the GJ are called connexin, and it is known that connexin holds transmembrane domains at four different sites and both of its N-terminal and C-terminal protrude on the cytoplasmic side of the membrane.

The TJ is an intercellular adhesion apparatus peculiar to epithelial cells and endothelial cells, where the cell membranes of contiguous cells are seen completely tightly apposed. The TJ surrounds individual cells and functions as a barrier to block or regulate permeation of water-soluble molecules between the luminal and basement membrane sides of a cell layer. It has also been described to act as a fence partitioning the cell membrane into apical and basolateral sides in order to maintain the polar distribution of such membrane proteins as ion channels and pumps as well as lipids on the cell membrane [Schneeberger, E. E. et al., "Am. J. Physiol.", Vol. 262, P. L647-L661, (1992)]. Owing to these functions of the TJ, milieus consisting of different fluid compositions are formed on the opposite sides of a cell layer, so that the polarity of the cell layer is maintained; hence the TJ can be said to be a fundamental structure of vital importance to multicellular organisms.

However, analysis of the molecular structure of the TJ has been less progressing, compared to other adhesion apparatus. In fact, it has constituted a serious drawback to the pursuit of molecular biological research on the TJ that the TJ adhesion molecule itself has not been identified yet.

The present inventors have established a method for isolation of AJ from rat liver, and have identified many proteins such as radixin and ZO-1 from this isolated AJ [Tsukita, Sh. et al., "Curr. Opin. Cell Biol.", Vol. 4, P. 834-839, (1992)]. From researches on ZO-1 and histologic findings for the AJ and the TJ, it can be presumed that the proteins in the AJ also contain protein of the TJ. In view of this, the present inventors have isolated AJ from chick liver, prepared a monoclonal antibody against the AJ as the antigen, and carried out structural analysis of the TJ-constituting protein using the antibody specifically reacting with the TJ. As a result, the present inventor has been successful in the structural analysis of a novel constituent protein dissimilar to known proteins, and designated the protein as occludin [Furuse, M. et al., "J. Cell Biol.", Vol. 123, p. 1777-1788, (1993)].

This chick occludin is a 56KDa protein composed of 504 amino acids, characterized conspicuously by transmembrane domains at four sites in the half of its N-terminus, with both the N- and C-terminals facing the cytoplasm and with two extracellular loops.

From subsequent studies, occludin was inferred to be an important factor in the analysis of physiological function of the TJ at the cellular level as well as at the whole body level, and drew much attention of investigators.

No further study has progressed, nevertheless, since the said protein has its origin in the chicken species greatly remote from humans. Thus, structural analysis of occludin of human origin has been expected for the sake of elucidation of the physiological function and medical analysis of the TJ. There is as yet no report of success in the elucidation of human occludin despite worldwide competition in research for this purpose in this field.

SUMMARY OF THE INVENTION

It is accordingly an object of the present invention to provide amino acid sequences of human, canine and mouse occludins and the DNAs encoding them.

In their report on the gene for human neuronal apoptosis inhibitory protein (NAIP), Roy et al. documented occurrence of a DNA fragment possessing a base sequence analogous to the C-terminal region of chick occludin in NAIP gene deletion mutants [Roy, N. et al., "Cell", Vol. 80, p. 167-178, (1995)]. To ascertain whether the said sequence actually encoded a part of human analogue of occludin or not, the present inventor selected primers out of the base sequence analogous to that of chick occludin and made a scrupulous screening with a cDNA library of human intestinal epithelial cell strain T84 as a template for PCR. The present inventor has thus succeeded in the analysis of the whole structure of human occludin. Further, the inventor has completed analyses of mouse and canine occludins, prepared anti-occludin monoclonal antibodies, and verified with histologic staining that the occludins were transmembrane type TJ proteins.

Best Mode for Carrying out the Invention:

The present invention is concerned with amino acid sequences of a human, a canine and a mouse occludin, DNAs for encoding them, anti-occludin antibodies, and a genetic analysis method utilizing them. More specifically, the present invention is directed to the following aspects.

(1) A DNA for encoding a human occludin having an amino acid sequence described in Sequence No. 1.
(2) A DNA for encoding a human occludin as described in Sequence No. 4.
(3) A DNA described in Sequence No. 5 for encoding a canine occludin having an amino acid sequence described in Sequence No. 2.
(4) A DNA described in Sequence No. 6 for encoding a mouse occludin having an amino acid sequence described in Sequence No. 3.
(5) A human, a canine and a mouse occludin having the amino acid sequences described in Sequence Nos. 1, 2 and 3, respectively.
(6) An occludin variant having an amino acid sequence in which one or plural amino acids in the amino acid sequence of each occludin are added, deleted or substituted, and a DNA for encoding the variant.
(7) A vector which comprises any of DNAs for encoding a human, a canine or a mouse occludin or for encoding their variants.
(8) A transformant which holds the vector.
(9) A method for manufacturing an occludin protein which comprises the steps of cultivating the transformant, and collecting an expressed product.
(10) A DNA probe comprising containing the whole or part of the base sequence as defined in Sequence No. 4, 5 or 6.
(11) A DNA primer comprising containing part of the base sequence as defined in Sequence No. 4, 5 or 6.
(12) A polyclonal antibody or a monoclonal antibody specifically binding to a human, a canine or a mouse occludin protein.
(13) An assay method and an assay reagent for occludin in a biological specimen, wherein an anti-occludin antibody is used.
(14) An analysis method of an occludin gene in a biological specimen, wherein said DNA primer or said DNA probe is used.
(15) A screening method of a drug affecting the expression of occludin, wherein occludin-expressing cells and an analyte are allowed to coexist, and an expression quantity of an occludin gene of said cells is then determined by the use of a DNA primer or a DNA probe.
(16) An antisense DNA derived from a human occludin DNA.
(17) A laboratory animal whose occludin DNA is knocked out.

Detailed Description of the Invention:

With the success of the present inventor in identifying occludin analogues of mammals, it is now possible to structurally and functionally test the constitution and function of the TJ at the molecular level. The barrier and fencing functions of the TJ and the related regulatory mechanisms can be analyzed through experiments involving control of expression of the gene for occludin or inhibition of the occludin function with either an antisense probe or an antibody, using various types of cultured human, mouse and canine (MDCK) cells. For example, it is now possible to determine whether or not overexpression of occludin cDNA gives rise to an increase in number of TJ strands seen in freeze-frac-

tured replicas and incidentally to an augmentation of the barrier function. Furthermore, the present invention has made it possible to establish a simple screening method for drugs affecting TJ function. For example, drugs affecting TJ function can be screened using various types of cells expressing occludin, by allowing the cell to react with a test article and subsequently by measuring the amount of cellular occludin gene or occludin protein expression. The gene analysis can be carried out by using a DNA probe or primer or other devices. It may be conducted by known methods, e.g. the Northern blot technique or Southern blot technique wherein RNA or DNA extracted from a test sample in the usual manner is pretreated when necessary, then electrophoresed on a membrane or gel, and hybridized with a labeled DNA probe, and the polymerase chain reaction (PCR) technique wherein the objective DNA is amplified using primers of about 20 bases corresponding to the relevant site and with a genomic DNA or cDNA as the template. The occludin protein can be quantitated, for example, by the use of an antibody.

Moreover, it is also made possible to ascertain how the TJ formation is involved in the morphogenesis of various organs and whether functional failure of the TJ has any relation to various pathologic states such as inflammation and tumor metastasis, by preparing various types of mutant mice and occludin gene-knocked out mice. The possibility of controlling a TJ function, especially its barrier function, is also of interest in connection with drug permeability. Thus, it would be feasible to control the blood-brain barrier via up- or down-regulation of occludin synthesis in epithelial cells of the brain. Control of the TJ function in the enteric epithelial cells is necessary to regulate drug absorption from the intestine. It will thus become possible to control drug absorption, particularly distribution to the brain tissue, by administering an effective substance screened out of drugs affecting the TJ function. Hence, the present invention is highly anticipated for elucidation of the physiological mechanisms primarily of the blood-brain barrier, as well as for analysis, diagnosis and treatment of disease states.

The DNA provided by the present invention can be utilized in the analysis of genes for occludin proteins and of gene expression thereof by using a part of it as a primer or a probe. The term a part here denotes that the oligonucleotide to be used as a primer or probe comprises containing at least a 10-relevant-base sequence, or preferably at least a 15-base sequence, or more preferably a corresponding polynucleotide comprising containing approximately 20- to 30-base sequence based on the DNA sequence of the present invention. As the probe, a higher macromolecular or even the whole DNA may be used.

There is a method utilizing antisense DNA or antisense RNA as a means to control the function of occludin. The method is intended to block the flow of gene expression by interfering with the reading of genetic information at any of the stages of gene expression such as DNA replication, transcription and translation, and the antisense technique employs nucleic acid or its analogue for the blockage (Wickstrome, E. ed., Prospects for Antisense Nucleic Acid Therapy of Cancer and AIDS. Wiley-Liss, New York, 1991). The elucidation of the occludin DNA according to the present invention has made possible the means for inhibiting occludin functions by the antisense method. The length of DNA oligomer has bearing on the double strand-forming capacity, membrane permeability and base sequence specificity; and at least 6 nucleotides, or preferably at least 10 mers, usually 15 to 30 mers, may be used. Appropriate sequences may be selected on the basis of the DNA sequence of the present invention, and verified by experimentation. Usually, the oligomer is chemically modified at its phosphate group, sugar moiety, and 3' and 5' tails in order to augment its stability (Cook, P.D., Anticancer Drig. Des., 5,585, 1991). Representative analogues are oligophosphorothioate where one of the oxygen atoms of the internucleoside phosphodiester group is replaced by a sulfur atom, and oligomethylsulfonate where the said oxygen is replaced by a methyl group; all such analogues are remarkably stable to nucleases. Besides, such oligomers as those with acridine or polylysine bonded to them and those containing N-methylthymidylate, to increase stability of the hybrid double strand, are also used. These oligomers can be synthesized by known chemical synthetic procedures. Antisense RNA derived from the DNA of the present invention may also be utilized.

The occludin protein of the present invention may be utilized for preparation of an antibody using the whole or a part of it as an epitope, and for use the antibody thus prepared as research and diagnostic reagents. The term epitope denotes an antigen determinant of polypeptide; epitope is usually comprised of at least 6 amino acids, and it is known that a polypeptide consisting of 6 amino acids combines with an antibody (JP-A-60-500684). The antigenic peptide of the subject protein signifies a polypeptide comprising a series of at least 6 amino acids, preferably a series of at least 8 amino acids, more preferably a series of at least 15 amino acids, or further preferably a series of at least 20 amino acids, based on the amino acid sequence of the present invention. Occludin provided by the present invention is a protein which, as inferred from its amino acid sequence in analogy with chick occludin, possesses transmembrane domains at four sites in a half of its N-terminal region, with the N- and C-terminus facing the cytoplasm, and which has two extracellular loops. In the case of human occludin of which amino acid 89-135 and 196-243 regions are presumed to be extracellularly apposed, various antibodies may be prepared by selecting antigenic sites appropriate for purposes and utilized as a means to elucidate the TJ function and as a means to suppression by the antibody of the TJ function. It is also possible to utilize the partial peptides as a means to screen compounds for those capable of binding to these peptides.

Proteins having an amino acid sequence of occludin of the present invention to which one or a plurality of amino acids are added or of which one or a plurality of amino acids are deleted or substituted are also encompassed by the

present invention.

(1) Preparation of cDNA library and structural analysis of occludin

Preparation of RNA may be carried out using human or animal cells (cell strains) as the raw material, by for example extraction with a mixed solution of guanidine thiocyanate, a surfactant, a chelating agent and a reductant, followed by phenol extraction, fractionation in organic solvents (Chamezynski et al., Anal. Biochem., 162, 156, 1987) and subsequently density gradient ultracentrifugation procedure. Using the RNA thus obtained as a template, a double strand DNA is prepared in the usual manner such as by the cDNA synthesis technique (Gubler, U. et al., Gene, 25, 263,1983) with the use of random primers, reverse transcriptase, DNA polymerase, etc. A DNA library can be prepared by insertion of the double strand DNA obtained into a bacteriophage such as λ zap or λ gt11 in the usual manner. Commercial cDNA libraries may also be used.

According to the report of Roy et al., it is possible to obtain a DNA fragment presumed to be of occludin DNA origin by properly selecting a primer region based on the base sequence analogous to the C-terminus of chick occludin, by amplification of the DNA with the PCR technique, and by its subclonation. Subsequent screening of the cDNA library with this DNA fragment as a probe and analysis of the base sequence of the clone isolated may yield a whole-length cDNA for occludin. The structure of the base sequence is determined by the Maxam-Gilbert method (Maxam, A. M. and Gilbert, W., Proc. Natl. Acad. Sci. USA, 74, 560, 1977) or by the dideoxynucleotide chain termination method (Sanger, F., Proc. Natl. Acad. Sci. USA, 74, 5463, 1977). The amino acid sequence is thus deduced on the basis of the base sequence. These gene manipulations can be performed by the known usual methods, for example in accordance with those described in Molecular Cloning. A Laboratory Manual, T. Maniatis et al. eds. (1989), Cold Spring Harbor Laboratory.

(2) Preparation of antibodies

To prepare the monoclonal antibody according to the present invention, human, canine or mouse occludin is used as the antigen, its complex with a carrier protein is prepared if deemed necessary, and appropriate animals are immunized by inoculation with the antigen. Antibody-forming cells obtained from the spleen or lymphnodes of the above immunized animals are fused with myeloma cells, whereby hybridomas producing an antibody strongly specific to occludin are selected to prepare the monoclonal antibody. The preparation procedure may be in accordance with the known prior method.

As the immunogen, any of such products as purified natural products and products prepared by genetic recombination technique or chemical synthesis may be used. For preparation of occludin by the recombinant DNA technique, the cDNA encoding occludin can be religated to the promoter downstream of a vector appropriate for expression of occludin by the known method using restriction enzymes and DNA ligase to obtain a recombinant expression vector. The said vector is nonlimitative insofar as it can be replicated and amplified in a host. With regard to the promoter and terminator, there is also no particular limitation as long as they are concordant with the host used for expression of the base sequence encoding occludin; and appropriate combinations suited to the host may also be practicable. The recombinant expression vector thus obtained is introduced into the host by the competent cell technique (J. Mol. Biol., 53, 154, 1970) or the calcium phosphate procedure (Science, 221, 551, 1983) to prepare a transformant. Such organisms as *Escherichia coli* and animals are used as the host, and the transformant obtained is cultured in an appropriate medium suited to the host. The culture incubation is carried out usually at a temperature between 20°C and 45°C and at pH between 5 and 8, with aeration and/or stirring where required. Isolation and purification of occludin from the cultured microorganisms or cells may be performed by an appropriate combination of known methods of isolation and purification. These known methods include salting out, organic solvent method, dialysis, gel filtration, electrophoresis, ion exchange chromatography, affinity chromatography, and reverse-phase high performance liquid chromatography.

The immunogen, occludin, is preferably to retain its whole structure but may be in the form of a fragment or peptide having its partial structure; it may be appropriately selected from the whole amino acid sequence of occludin. For preparation of the fragment or peptide, a method such as chemical synthesis, the above mentioned gene recombination procedure or degradation of a naturally occurring article is employed.

Various condensing agents may be used for preparation of the immunogen-carrier protein complex; such reagents as glutaraldehyde, carbodiimide, and maleimide activated ester may be used.

The carrier protein may be any of those commonly used such as bovine serum albumin, thyroglobulin and hemocyanin, and usually the method wherein a 1- to 5-fold quantity of a carrier protein is coupled to antigen is used.

Animals employed for immunization include the mouse, rat, rabbit, and guinea pigs, and inoculation is made by subcutaneous, intramuscular or intraperitoneal injection. The administration of immunogen may be carried out in the form of a mixture with complete Freund adjuvant or with incomplete adjuvant, and is usually made once every 2 to 5 weeks. Antibody-producing cells obtained from the spleen or lymphnodes of the immunized animals are fused with myeloma

cells and isolated as hybrydomas. The myeloma cells used are those of mouse, rat or human origin, preferably allogeneic to the antigen-producing cells used but, in some instances, can be xenogeneic.

The manipulation of cell fusion can be conducted in accordance, for example, with the method of Milstein and Köhler (Nature, 256,495, 1975). Fusogens used include such agents as polyethylene glycol and Sendai virus, and the cell fusion can be made by incubating antibody-producing cells with myeloma cells in an approximate population ratio of 1:1 to 10:1 at a temperature between 20 and 40°C, preferably between 30 and 37°C, for about 1 to 10 minutes, using polyethylene glycol (mean molecular weight: 1,000 to 4,000) usually at a concentration of about 20 to 50%.

Various immunochemical methods can be used for screening hybridomas producing an anti-occludin antibody. The methods include enzyme-linked immunosorbent assay (ELISA) using microplates coated with occludin, enzyme immunoassay (EIA) using micoplates coated with an anti-immunoglobulin antibody, and Western blotting technique in which samples containing occludin are electrophoresed with the subsequent use of nitrocellulose transfer membranes.

Clones are obtained from these wells further by, for example, limiting dilution. Screening and breeding of hybridomas are performed in a culture medium for animal cells (e.g. RPMI 1640) containing 10-20% fetal bovine serum usually with added HAT (hypoxanthine, aminopterin and thymidine). The clone thus obtained is intraperitoneally transplanted into BALB/c mice previously dosed with bristan, and ascites containing a high concentration of a monoclonal antibody is collected 10-14 days later so that the ascites can be used as a source for purification of the antibody. Furthermore, the cloned hybridoma cells are cultured so that the cultured cells can be used as the source for purification of the antibody. Known methods for purification of immunoglobulin may be used for recovering the monoclonal antibody; the recovery can be readily accomplished for example by such means as ammonium sulfate fractionation, PEG fractionation, ethanol fractionation, utilization of anion exchangers, and affinity chromatography.

With immunological methods using the anti-occludin monoclonal antibody obtained in accordance with the present invention, it is possible to make qualitative and quantitative determination of occludin in biological specimens. As the immunological methods, conventional methods such as immunohistologic staining, enzyme immunoassay, agglutination test, competitive assay, and sandwich technique may be applied to samples from biological specimens that have been appropriately processed, where required, e.g. isolation of cells and extraction. The immunohistologic staining can be performed for example by the direct method using a labeled antibody or the indirect method using a labeled antibody directed to the antibody bound to target antigen. Any of such known labeling substances as fluorescent agents, radioactive substances, enzymes, metals and dyes may be used as labeling agents.

The monoclonal antibody of the present invention may be used in the form of Fab' or Fab fragment after removal of its Fc' or Fc region, or in the form of polymer of either fragment. Furthermore, it may also be in the form of a chimeric antibody, a humanized antibody, or a human antibody.

EXAMPLE

The present invention will now be illustrated in detail with specific embodiments by the following examples. Of course, the present invention shall not be limited to the following examples.

Example 1 Structural analysis of human occludin

Based on a base sequence of a part of human NAIP-deficient gene analogous to the C-terminus of chick occludin, PCR was performed using as primers the oligonucleotides of Sequence Nos. 7 and 8. λgt11 cDNA library was prepared by purifying poly(A)+RNA from a source consisting of human intestinal epithelial cell strain T84 and by using TimeSaver cDNA Synthesis Kit (trade name; Pharmacia LKB Biotechnology Inc.) and GIGAPACK II Packaging Extract (Stratagene Inc.). The PCR was carried out with this library as the template and with said two primers yielded a cDNA fragment of 363 base pairs.

This DNA fragment was DIG-labeled using DIG Labeling Kit (trade name; Boehringer Mannheim), and said library was screened with it used as the probe. As a result, three cDNA clones were isolated, their insertion sites were cut, and they were subcloned to pBluescript SK(+). Of these, the two clones phOc6 and phOc16 were presumed to contain a total ORF, and these two cloned strands were analyzed for their base sequences, with the results demonstrating that said base sequences encoded the whole structure of human occludin. The coding sequence was determined using a 7-deaza Sequenase Version Deoxy Terminator Cycle Sequencing Kit (trade name; Applied Bicsystems). The base sequence is shown in Sequence No. 4, and the amino acid sequence deduced therefrom in Sequence No. 1.

Structure of canine and mouse occludins were determined in the same manner as described above using λgt11 and λgt10 cDNA libraries, respectively, prepared from dog kidney (MDCK) cells and mouse lung cells. The base sequence and amino acid sequence of canine occludin are shown in Sequence Nos. 5 and 2, and the base sequence and amino acid sequence of mouse occludin in Sequence Nos. 6 and 3. *Escherichia coli* JM 109 containing the human occludin cDNA has been deposited (Deposition No. FERM BP-5477) with the National Institute of Bioscience and Human Technology, Ministry of International Trade and Industry, Japan (address: 1-1-3, Higashi, Tsukuba, Ibaraki, 305

Japan) as of March 15, 1996.

Example 2 Preparation of an anti-human occludin monoclonal antibody

The cDNA fragment encoding the cytoplasmic region on the C-terminal side of human occludin was obtained by cutting the pBluescript SK(-) vector containing Sequence No. 4 described in Example 1 with restriction enzymes Ssp I and EcoR I (both being products of Takara Shuzo Co., Ltd.). This fragment was introduced into pGEX-3X vector, and *Escherichia coli* transformed with this vector was cultured to prepare a GST fused protein. Rats were immunized with this fused protein as an antigen, so that a monoclonal antibody was prepared.

The rat immunization was performed by injecting the antigen in doses of 300 μg/injection into the hindlimb paw, first as an emulsion with complete Freund adjuvant and the antigen alone twice thereafter (days 3 and 7 after the first). On the day following the last injection, inguinal lymphnodes were excised from the immunized animals and used for cellular fusion.

The rat lymphocytes and mouse myeloma P3 cells were combined in a ratio of 2.5:1, and the mixture was incubated in RPMI medium containing 1 g of polyethylene glycol (mean MW 4,000) dissolved in it, for 2 minutes according to a modified method of Köhler et al., to permit fusion of the cells. Fused cells were seeded in 24-well plates with HAT medium containing 10% HCF (Bokusui-Braun) for 9 days, followed by incubation in HT medium and subsequently in flasks with RPMI medium. Hybridomas were cloned by assaying supernatants of wells showing cellular growth for antibody titer using immunoblot technique and fluorescent antibody staining in cultures of human intestinal epithelial cell strain T84, and by limiting dilution from proper wells. The hybridoma cells were seeded at calculated concentration of 7 cells/well in microtiter plates, and screened by immunoblotting technique to verify and isolate clonal hybridoma cell strains. Antibodies were purified from culture supernatants of said hybridoma.

Example 3 Cell staining

Human intestinal epithelial cells were fixed in 3% formalin in phosphate buffered saline (PBS) at room temperature for 15 minutes, and further treated with 0.2% Triton X-100 in PBS at room temperature for 15 minutes. After blocking the cells with 1% bovine serum albumin (BSA), the test substance was added and incubated for 30 minutes at room temperature. After subsequent washing, FITC-labeled anti-rat immunoglobulin antibody was added and incubated for 30 minutes at room temperature, followed by washing off unreacted antibody and examination with a fluorescent microscope.

Results of double immunofluorescent staining with monoclonal antibody to the TJ-related protein ZO-1 and the anti-human occludin monoclonal antibody of the present invention are shown in the figure. As totally the same staining pattern as that of ZO-1 (reported in the literature) was observed, the human protein of the present invention has proven to be a human homologue of the TJ adhesion molecule occludin.

Example 4 Expression of occludin in cerebrovascular cells

Since cerebral vascular endothelial cells are thought to have a high electroresistant TJ, which form the brain-blood barrier unlike peripheral vascular endothelial cells, I examined the distribution and expression of occludin in cultured porcine brain vascular endothelial cells (PBEC) possessing the high electroresistant TJ and cultured porcine aortic endothelial cells (PAEC).

As porcine occludin cDNA fragment, a 363 base fragment was prepared by amplification with PCR technique using as primers 1359-1391 sense strand (Sequence No. 7) and 1692-1721 antisense strand (Sequence No. 8) from the human occludin DNA sequence (Sequence No. 4). The amino acid sequence based on analysis of the coding base sequence of said fragment showed a high degree of homology with amino acid sequences of human, mouse and canine occludins, thus verifying the fragment to be a cDNA for porcine occludin. ^{32}P-labeled said fragment was used as a probe.

To prepare mRNA from the cultured cells, an agarose gel electrophoresed sample was transferred onto nitrocellulose membrane and hybridized with the c DNA probe under highly stringent conditions, using an RNA isolation kit (Stratagene). As a result, the occludin mRNA showed a strong band at about 2.4 kb in PBEC, whereas in PAEC, only a very weak band was noted at that position.

Expression of occludin in these cells was compared using anti-mouse occludin antibody as a monoclonal antibody specifically recognizing mammalian occludins and an antibody against the TJ-related protein ZO-1.

Anti-mouse occludin rat antibody was prepared using mouse occludin:glutathione-S-transferase fused protein as a antigen, and FITC-labeled anti-rat IgG sheep antibody was used for detection of said antibody. When equal protein quantities of extracts from disrupted cultured cells were analyzed by immunoblotting after one-dimensional gel electrophoresis, a strong band of occludin was detected at about 58KD in PBEC while a considerably weaker band was noted

at that position in PAEC. On the other hand, there was no appreciable difference in expression of ZO-1 between the two types of cells. With immunostaining, PBEC exhibited a marked occludin expression with the same continuous intercellular localization as ZO-1, as seen in the immunoblotting study. In PAEC, in contrast, occludin was scarcely detected and ZO-1 showed a discontinuous intercellular localization. These results suggest that the relatively marked expression of occludin in PBEC is required for the formation of the highly electroresistant TJ, and provide evidence that occludin is the constituent protein of the TJ.

SEQUENCE

SEQ ID NO: 1

SEQUENCE LENGTH: 522

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE

ORGANISM: human (intestinal epithelial cell strain T84)

SEQUENCE DESCRIPTION

Met Ser Ser Arg Pro Leu Glu Ser Pro Pro Pro Tyr Arg Pro Asp Glu
1 5 10 15

Phe Lys Pro Asn His Tyr Ala Pro Ser Asn Asp Ile Tyr Gly Gly Glu
20 25 30

Met His Val Arg Pro Met Leu Ser Gln Pro Ala Tyr Ser Phe Tyr Pro
35 40 45

Glu Asp Glu Ile Leu His Phe Tyr Lys Trp Thr Ser Pro Pro Gly Val
50 55 60

Ile Arg Ile Leu Ser Met Leu Ile Ile Val Met Cys Ile Ala Ile Phe
65 70 75 80

Ala Cys Val Ala Ser Thr Leu Ala Trp Asp Arg Gly Tyr Gly Thr Ser
85 90 95

Leu Leu Gly Gly Ser Val Gly Tyr Pro Tyr Gly Gly Ser Gly Phe Gly
100 105 110

Ser Tyr Gly Ser Gly Tyr Gly Tyr Gly Tyr Gly Tyr Gly Tyr Gly Tyr
115 120 125

Gly Gly Tyr Thr Asp Pro Arg Ala Ala Lys Gly Phe Met Leu Ala Met

130 135 140

Ala Ala Phe Cys Phe Ile Ala Ala Leu Val Ile Phe Val Thr Ser Val

145 150 155 160

Ile Arg Ser Glu Met Ser Arg Thr Arg Arg Tyr Tyr Leu Ser Val Ile

165 170 175

Ile Val Ser Ala Ile Leu Gly Ile Met Val Phe Ile Ala Thr Ile Val

180 185 190

Tyr Ile Met Gly Val Asn Pro Thr Ala Gln Ser Ser Gly Ser Leu Tyr

195 200 205

Gly Ser Gln Ile Tyr Ala Leu Cys Asn Gln Phe Tyr Thr Pro Ala Ala

210 215 220

Thr Gly Leu Tyr Val Asp Gln Tyr Leu Tyr His Tyr Cys Val Val Asp

225 230 235 240

Pro Gln Glu Ala Ile Ala Ile Val Leu Gly Phe Met Ile Ile Val Ala

245 250 255

Phe Ala Leu Ile Ile Phe Phe Ala Val Lys Thr Arg Arg Lys Met Asp

260 265 270

Arg Tyr Asp Lys Ser Asn Ile Leu Trp Asp Lys Glu His Ile Tyr Asp

275 280 285

Glu Gln Pro Pro Asn Val Glu Glu Trp Val Lys Asn Val Ser Ala Gly

290 295 300

Thr Gln Asp Val Pro Ser Pro Pro Ser Asp Tyr Val Glu Arg Val Asp

305 310 315 320

Ser Pro Met Ala Tyr Ser Ser Asn Gly Lys Val Asn Asp Lys Arg Phe

325 330 335

Tyr Pro Glu Ser Ser Tyr Lys Ser Thr Pro Val Pro Glu Val Val Gln

340 345 350

Glu Leu Pro Leu Thr Ser Pro Val Asp Asp Phe Arg Gln Pro Arg Tyr

355 360 365

Ser Ser Gly Gly Asn Phe Glu Thr Pro Ser Lys Arg Ala Pro Ala Lys

370 375 380

Gly Arg Ala Gly Arg Ser Lys Arg Thr Glu Gln Asp His Tyr Glu Thr

385 390 395 400

Asp Tyr Thr Thr Gly Gly Glu Ser Cys Asp Glu Leu Glu Glu Asp Trp

405 410 415

Ile Arg Glu Tyr Pro Pro Ile Thr Ser Asp Gln Gln Arg Gln Leu Tyr

420 425 430

Lys Arg Asn Phe Asp Thr Gly Leu Gln Glu Tyr Lys Ser Leu Gln Ser

435 440 445

Glu Leu Asp Glu Ile Asn Lys Glu Leu Ser Arg Leu Asp Lys Glu Leu

450 455 460

Asp Asp Tyr Arg Glu Glu Ser Glu Glu Tyr Met Ala Ala Ala Asp Glu

465 470 475 480

Tyr Asn Arg Leu Lys Gln Val Lys Gly Ser Ala Asp Tyr Lys Ser Lys

485 490 495

Lys Asn His Cys Lys Gln Leu Lys Ser Lys Leu Ser His Ile Lys Lys

500 505 510

Met Val Gly Asp Tyr Asp Arg Gln Lys Thr

515 520

SEQUENCE ID NO: 2

SEQUENCE LENGTH: 521

SEQ TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE

ORGANISM: canine (kidney cell MDCK)

SEQUENCE DESCRIPTION

```
Met Ser Ser Arg Pro Phe Glu Ser Pro Pro Pro Tyr Arg Pro Asp Glu
 1               5                  10                  15
Phe Lys Pro Asn His Tyr Ala Pro Ser Asn Asp Val Tyr Gly Gly Asp
            20                  25                  30
Met His Val Arg Pro Met Leu Ser Gln Pro Ala Tyr Ser Phe Tyr Pro
        35                  40                  45
Glu Asp Glu Ile Leu His Phe Tyr Lys Trp Thr Ser Pro Pro Gly Val
    50                  55                  60
Ile Arg Ile Leu Ser Met Leu Val Ile Val Met Cys Ile Ala Ile Phe
65                  70                  75                  80
Gly Cys Val Ala Ser Thr Leu Ala Trp Asp Arg Gly Tyr Gly Thr Gly
                85                  90                  95
Leu Met Gly Gly Ser Ile Gly Tyr Pro Tyr Gly Ser Gly Phe Gly Ser
            100                 105                 110
Tyr Gly Thr Gly Tyr Gly Tyr Gly Phe Gly Tyr Gly Tyr Gly Tyr Gly
        115                 120                 125
Gly Tyr Thr Asp Pro Arg Ala Ala Lys Gly Phe Leu Leu Ala Met Val
    130                 135                 140
Ala Phe Cys Phe Ile Ala Ala Leu Val Ile Phe Val Thr Ser Val Ile
145                 150                 155                 160
```

Arg Ser Asp Ile Ser Arg Thr Arg Arg Tyr Tyr Leu Thr Val Ile Ile
165 170 175
180 185 190
Ile Met Gly Val Asn Pro Thr Ala Gln Ala Ser Gly Ser Leu Tyr Ser
195 200 205
Ser Gln Ile Tyr Ala Met Cys Asn Gln Phe Tyr Ala Ser Thr Ala Thr
210 215 220
Gly Leu Tyr Met Asp Gln Tyr Leu Tyr His Tyr Cys Val Val Asp Pro
225 230 235 240
Gln Glu Ala Ile Ala Ile Val Leu Gly Phe Met Val Ile Val Ala Phe
245 250 255
Ala Leu Ile Ile Phe Phe Ala Val Lys Thr Arg Arg Lys Met Asp Arg
260 265 270
Tyr Asp Lys Ser Asn Ile Leu Trp Asp Lys Glu His Ile Tyr Asp Glu
275 280 285
Gln Pro Pro Asn Val Glu Glu Trp Val Lys Asn Val Ser Ala Gly Thr
290 295 300
Gln Asp Met Pro Pro Pro Pro Ser Asp Tyr Val Glu Arg Val Asp Ser
305 310 315 325
Pro Met Ala Tyr Ser Ser Asn Gly Lys Val Asn Asp Lys Arg Leu Tyr
325 330 335
Pro Glu Ser Ser Tyr Lys Ser Thr Pro Val Pro Glu Val Val Gln Glu
340 345 350
Leu Pro Ala Thr Ser Pro Ala Asp Asp Phe Arg Gln Pro Arg Tyr Ser
355 360 365
Ser Ser Gly His Leu Glu Pro Pro Ser Lys Arg Ala Pro Ser Lys Gly

370 375 380

Arg Thr Gly Arg Pro Lys Arg Leu Glu Gln Asp His Tyr Glu Thr Asp

385 390 395 400

Tyr Thr Thr Gly Gly Glu Ser Cys Asp Glu Leu Glu Glu Asp Trp Ile

405 410 415

Arg Glu Tyr Pro Pro Ile Thr Ser Asp Gln Gln Arg Gln Leu Tyr Lys

420 425 430

Arg Asn Phe Asp Thr Gly Leu Gln Glu Tyr Lys Ser Leu Gln Ala Glu

435 440 445

Leu Asp Glu Ile Asn Lys Glu Leu Ser Arg Leu Asp Lys Glu Leu Asp

450 455 460

Asp Tyr Arg Glu Glu Ser Glu Glu Tyr Met Ala Ala Ala Asp Glu Tyr

465 470 475 480

Asn Arg Leu Lys Gln Val Lys Gly Ser Pro Asp Tyr Lys Asn Lys Arg

485 490 495

Asn Tyr Cys Lys Gln Leu Lys Ser Lys Leu Ser His Ile Lys Lys Met

500 505 510

Val Gly Asp Tyr Asp Arg Gln Lys Thr

515 520

SEQ ID NO: 3

SEQUENCE LENGTH: 521

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE

ORGANISM: mouse (lung cell)

SEQUENCE DESCRIPTION

Met Ser Val Arg Pro Phe Glu Ser Pro Pro Pro Tyr Arg Pro Asp Glu
1 5 10 15

Phe Lys Pro Asn His Tyr Ala Pro Ser Asn Asp Met Tyr Gly Gly Glu
20 25 30

Met His Val Arg Pro Met Leu Ser Gln Pro Ala Tyr Ser Phe Tyr Pro
35 40 45

Glu Asp Glu Ile Leu His Phe Tyr Lys Trp Thr Ser Pro Pro Gly Val
50 55 60

Ile Arg Ile Leu Ser Met Leu Ile Ile Val Met Cys Ile Ala Ile Phe
65 70 75 80

Ala Cys Val Ala Ser Thr Leu Ala Trp Asp Arg Gly Tyr Gly Thr Gly
85 90 95

Leu Phe Gly Gly Ser Leu Asn Tyr Pro Tyr Ser Gly Phe Gly Tyr Gly
100 105 110

Gly Gly Tyr Gly Gly Gly Tyr Gly Gly Tyr Gly Tyr Gly Tyr Gly Gly
115 120 125

Tyr Thr Asp Pro Arg Ala Ala Lys Gly Phe Leu Leu Ala Met Ala Ala
130 135 140

Phe Cys Phe Ile Ala Ser Leu Val Ile Phe Val Thr Ser Val Ile Arg
145 150 155 160

Ser Gly Met Ser Arg Thr Arg Arg Tyr Tyr Leu Ile Val Ile Ile Val
165 170 175

Ser Ala Ile Leu Gly Ile Met Val Phe Ile Ala Thr Ile Val Tyr Ile
180 185 190

Met Gly Val Asn Pro Thr Ala Gln Ala Ser Gly Ser Met Tyr Gly Ser
195 200 205
Gln Ile Tyr Met Ile Cys Asn Gln Phe Tyr Thr Pro Gly Gly Thr Gly
210 215 220
Leu Tyr Val Asp Gln Tyr Leu Tyr His Tyr Cys Val Val Asp Pro Gln
225 230 235 240
Glu Ala Ile Ala Ile Val Leu Gly Phe Met Ile Ile Val Ala Phe Ala
245 250 255
Leu Ile Ile Phe Phe Ala Val Lys Thr Arg Arg Lys Met Asp Arg Tyr
260 265 270
Asp Lys Ser Asn Ile Leu Trp Asp Lys Glu His Ile Tyr Asp Glu Gln
275 280 285
Pro Pro Asn Val Glu Glu Trp Val Lys Asn Val Ser Ala Gly Thr Gln
290 295 300
Asp Met Pro Pro Pro Pro Ser Asp Tyr Ala Glu Arg Val Asp Ser Pro
305 310 315 320
Met Ala Tyr Ser Ser Asn Gly Lys Val Asn Gly Lys Arg Ser Tyr Pro
325 330 335
Glu Ser Phe Tyr Lys Ser Thr Pro Leu Val Pro Glu Val Ala Gln Glu
340 345 350
Ile Pro Leu Thr Leu Ser Val Asp Asp Phe Arg Gln Pro Arg Tyr Ser
355 360 365
Ser Asn Gly Asn Leu Glu Thr Pro Ser Lys Arg Ala Pro Thr Lys Gly
370 375 380
Lys Ala Gly Lys Gly Lys Arg Thr Asp Pro Asp His Tyr Glu Thr Asp
385 390 395 400

Tyr Thr Thr Gly Gly Glu Ser Cys Glu Glu Leu Glu Glu Asp Trp Val
405 410 415
Arg Glu Tyr Pro Pro Ile Thr Ser Asp Gln Gln Arg Gln Leu Tyr Lys
420 425 430
Arg Asn Phe Asp Ala Gly Leu Gln Glu Tyr Lys Ser Leu Gln Ala Glu
435 440 445
Leu Asp Asp Val Asn Lys Glu Leu Ser Arg Leu Asp Lys Glu Leu Asp
450 455 460
Asp Tyr Arg Glu Glu Ser Glu Glu Tyr Met Ala Ala Ala Asp Glu Tyr
465 470 475 480
Asn Arg Leu Lys Gln Val Lys Gly Ser Ala Asp Tyr Lys Ser Lys Arg
485 490 495
Asn Tyr Cys Lys Gln Leu Lys Ser Lys Leu Ser His Ile Lys Arg Met
500 505 510
Val Gly Asp Tyr Asp Arg Arg Lys Pro
515 520

SEQ ID NO: 4

SEQUENCE LENGTH: 2379

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE

ORGANISM: human (intestinal epithelial cell strain T34)

FEATURE

Feature Key: mat peptide

Feature Table Definition: 168.. 1733

SEQUENCE DESCRIPTION

CTCCCGCGTC CACCTCTCCC TCCCTGCTTC CTCTGGCGGA GGCGGCAGGA ACCGAGAGCC 60

AGGTCCAGAG CGCCGAGGAG CCGGTCTAGG ACGCAGCAGA TTGGTTTATC TTGGAAGCTA 120

AAGGGCATTG CTCATCCTGA AGATCAGCTG ACCATTGACA ATCAGCCATG TCATCCAGCC 180

CTCTTGAAAG TCCACCTCCT TACAGGCCTG ATGAATTCAA ACCGAATCAT TATGCACCAA 240

GCAATGACAT ATATGGTGGA GAGATGCATG TTCGACCAAT GCTCTCTCAG CCAGCCTACT 300

CTTTTTACCC AGAAGATGAA ATTCTTCACT TCTACAAATG GACCTCTCCT CCAGGAGTGA 360

TTCGGATCCT GTCTATGCTC ATTATTGTGA TGTGCATTGC CATCTTTGCC TGTGTGGCCT 420

CCACGCTTGC CTGGGACAGA GGCTATGGAA CTTCCCTTTT AGGAGGTAGT GTAGGCTACC 480

CTTATGGAGG AAGTGGCTTT GGTAGCTACG GAAGTGGCTA TGGCTATGGC TATGGTTATG 540

GCTATGGCTA CGGAGGCTAT ACAGACCCAA GAGCAGCAAA GGGCTTCATG TTGGCCATGG 600

CTGCCTTTTG TTTCATTGCC GCGTTGGTGA TCTTTGTTAC CAGTGTTATA AGATCTGAAA 660

TGTCCAGAAC AAGAAGATAC TACTTAAGTG TGATAATAGT GAGTGCTATC CTGGGCATCA 720

TGGTGTTTAT TGCCACAATT GTCTATATAA TGGGAGTGAA CCCAACTGCT CAGTCTTCTG 780

GATCTCTATA TGGTTCACAA ATATATGCCC TCTGCAACCA ATTTTATACA CCTGCAGCTA 840

CTGGACTCTA CGTGGATCAG TATTTGTATC ACTACTGTGT TGTGGATCCC CAGGAGGCCA 900

TTGCCATTGT ACTGGGGTTC ATGATTATTG TGGCTTTTGC TTTAATAATT TTCTTTGCTG 960

TGAAAACTCG AAGAAAGATG GACAGGTATG ACAAGTCCAA TATTTTGTGG GACAAGGAAC 1020

ACATTTATGA TGAGCAGCCC CCCAATGTCG AGGAGTGGGT TAAAAATGTG TCTGCAGGCA 1080

CACAGGACGT GCCTTCACCC CCATCTGACT ATGTGGAAAG AGTTGACAGT CCCATGGCAT 1140

ACTCTTCCAA TGGCAAAGTG AATGACAAGC GGTTTTATCC AGAGTCTTCC TATAAATCCA 1200

CGCCGGTTCC TGAAGTGGTT CAGGAGCTTC CATTAACTTC GCCTGTGGAT GACTTCAGGC 1260

AGCCTCGTTA CAGCAGCGGT GGTAACTTTG AGACACCTTC AAAAAGAGCA CCTGCAAAGG 1320

GAAGAGCAGG AAGGTCAAAG AGAACAGAGC AAGATCACTA TGAGACAGAC TACACAACTG 1380

GCGGCGAGTC CTGTGATGAG CTGGAGGAGG ACTGGATCAG GGAATATCCA CCTATCACTT 1440
CAGATCAACA AAGACAACTG TACAAGAGGA ATTTTGACAC TGGCCTACAG GAATACAAGA 1500
GCTTACAATC AGAACTTGAT GAGATCAATA AAGAACTCTC CCGTTTGGAT AAAGAATTGG 1560
ATGACTATAG AGAAGAAAGT GAAGAGTACA TGGCTGCTGC TGATGAATAC AATAGACTCA 1620
AGCAAGTGAA GGGATCTGCA GATTACAAAA GTAAGAAGAA TCATTGCAAG CAGTTAAACA 1680
GCAAATTGTC ACACATCAAG AAGATGGTTG GAGACTATGA TAGACAGAAA ACATAGAAGG 1740
CTGATGCCAA GTTGTTTGAG AAATTAAGTA TCTGACATCT CTGCAATCTT CTCAGAAGGC 1800
AAATGACTTT GGACCATAAC CCCGGAAGCC AAACCTCTGT GAGCATCACA AAGTTTTGGT 1860
TGCTTTAACA TCATCAGTAT TGAAGCATTT TATAAATCGC TTTTGATAAT CAACTGGGCT 1920
GAACACTCCA ATTAAGGATT TTATGCTTTA AACATTGGTT CTTGTATTAA GAATGAAATA 1980
CTGTTTGAGG TTTTTAAGCC TTAAAGGAAG GTTCTGGTGT GAACTAAACT TTCACACCCC 2040
AGACGATGTC TTCATACCTA CATGTATTTG TTTGCATAGG TGATCTCATT TAATCCTCTC 2100
AACCACCTTT CAGATAACTG TTATTTATAA TCACTTTTTT CCACATAAGG AAACTGGGTT 2160
CCTGCAATGA AGTCTCTGAA GTGAAACTGC TTGTTTCCTA GCACACACTT TTGGTTAAGT 2220
CTGTTTTATG ACTTCATTAA TAATAAATTC CCTGGCCTTT CATATTTTAG CTACTATATA 2280
TGTGATGATC TACCAGCCTC CCTATTTTTT TTCTGTTATA TAAATGGTTA AAAGAGGTTC 2340
TTCTTAAATA ATAAAGATCA TGTAAAAGTA AAAAAAAAA 2379

SEQ ID NO: 5
SEQUENCE LENGTH: 1961
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE
ORGANISM: canine (kidney cell strain MDCK)

FEATURE

Feature Key: mat peptide

Feature Table Definition: 72.. 1624

SEQUENCE DESCRIPTION

CAGGTTGGCT TATTTTGGGG AGCTCTGGGA TCCTGCTCGT CCTGAAGATC GGGTGATCAT 60

TGACATCAGC CATGTCATCG AGGCCTTTTG AGAGTCCACC TCCGTATAGA CCTGATGAAT 120

TCAAACCCAA TCATTATGCA CCGAGCAATG ATGTGTACGG TGGGGACATG CACGTCCGAC 180

CCATGCTCTC TCAGCCGGCG TATTCTTTCT ACCCAGAAGA TGAAATTCTT CACTTCTACA 240

AATGGACCTC TCCTCCAGGA GTAATTCGGA TTCTGTCCAT GCTTGTCATT GTGATGTGCA 300

TCGCCATATT TGGCTGTGTC GCGTCCACGC TCGCCTGGGA TAGAGGCTAT GGAACTGGCT 360

TAATGGGTGG TAGCATAGGC TACCCTTACG GAAGTGGCTT CGGGAGCTAC GGGACTGGCT 420

ACGGCTACGG GTTTGGCTAC GGCTACGGCT ACGGCGGCTA CACGGATCCC AGAGCAGCAA 480

AGGGCTTCCT CCTGGCCATG GTGGCCTTTT GTTTTATCGC TGCATTGGTG ATATTTGTTA 540

CCAGCGTTAT AAGGTCTGAC ATATCCAGAA CCAGAAGGTA CTACTTGACT GTAATAATAC 600

TGAGTGCCTT CCTGGGCGTC ATGATGTTCA TTGCTACAAT TGTCTATATA ATGGGAGTCA 660

ATCCAACTGC CCAGGCTTCT GGGTCTTTAT ACAGTTCACA GATATATGCC ATGTGCAACC 720

AGTTCTATGC ATCTACAGCT ACCGGACTCT ACATGGATCA GTATTTGTAT CACTACTGTC 780

TGGTGGATCC CCAAGAGGCA ATTGCCATTG TCCTGGGATT CATGGTGATT GTGGCTTTTG 840

CTTTAATAAT TTTCTTTGCT GTGAAAACTC GAAGAAAGAT GGACCGGTAT GACAAGTCGA 900

ATATATTGTG GGACAAGGAA CATATTTATG ATGAACAACC CCCCAATGTT GAAGAGTGGG 960

TTAAAAACGT TTCTGCAGGC ACACAAGACA TGCCTCCTCC CCCTTCTGAC TATGTGGAGA 1020

GAGTGGACAG TCCCATGGCG TACTCTTCCA ATGGTAAAGT GAATGACAAG CGGTTGTATC 1080

CAGAGTCTTC CTATAAATCA ACACCGGTCC CCGAAGTGGT GCAGGAGCTG CCCGCCACCT 1140

CCCCTGCGGA TGACTTCAGG CAGCCTCGCT ACAGCAGCAG CGGGCACTTG GAGCCACCTT 1200

CGAAGAGGGC CCCCTCGAAA GGAAGAACGG GAAGGCCCAA GAGGCTGGAG CAGGACCACT 1260

ATGAGACAGA CTACACGACG GGCGGCGAGT CGTGTGACGA GCTGGAGGAG GACTGGATCA 1320

GGGAATATCC ACCTATCACT TCAGATCAAC AAAGACAACT CTACAAGAGA AATTTTGACA 1380

CTGGCCTGCA GGAATACAAG AGCTTACAAG CAGAACTTGA TGAGATCAAT AAAGAACTCT 1440

CTCGCCTGGA TAAAGAATTG GATGACTATA GAGAAGAAAG TGAAGAGTAC ATGGCTGCTG 1500

CTGATGAGTA CAATAGACTG AAGCAAGTTA AGGGATCTCC AGATTACAAA AATAAGAGGA 1560

ATTATTGCAA GCAGTTGAAG AGCAAATTGT CCCACATCAA GAAGATGGTT GGAGACTATG 1620

ATAGACAGAA AACATAGAAG GCAGATGCCA CACAGTTTGA GAGATTGTGA AGTATTTGAC 1680

ATATCTGCAA CGTTGTCAGA AGGCAGAATG ACTTTGGATT TCGAACCCAG GAGGCCAGAT 1740

CTTTGTGATC ATTACAAAGT TTTGGTAGCT TTAATATCAT CAGTATTGAA GCATTTTACA 1800

CATAGCTTTT GATAATCAAC TGGGCTGAAC ACTCCCGATT AAGGATTCTG TGCTTTAGAC 1860

TTTGGCTGTT GTGCTAAAGG ACTGAGTATA GGTGGAGGTT TTCAGACCTT GGAAGAAGGT 1920

CCCACGGTGA ACTTGTGCTG TGAACTTGCA CACTTGGGGC A 1961

SEQ ID NO: 6

SEQUENCE LENGTH: 2839

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE

ORGANISM: mouse (lung cell)

FEATURE

Feature Key: mat peptide

Feature Table Definition: 223..1785

SEQUENCE DESCRIPTION

GGAGTTTCAG GTGAATGGGT CACCGAGGGA GGAGGCTGGC CACGCCACAC CTCGTCGCTA 60

GTGCCCACCT CCCGGCCCCT CTTTCCTTAG GCGACAGCGG TGGAGTTGCG GGAGAGCGGT 120

CCAGCGCACG GAGCAACCGG CTAGGGGCTC GGCAGGTTCG CTTATCTTGG GAGCCTGGAC 180
ATTTTGCTCA TCATAAAGAT TAGGTGACCA GTGACATCAG CCATGTCCGT GAGGCCTTTT 240
GAAAGTCCAC CTCCTTACAG ACCTGATGAA TTCAAACCCA ATCATTATGC ACCAAGCAAT 300
GACATGTATG GCGGAGAGAT GCATGTCCGG CCGATGCTCT CTCAGCCAGC GTACTCTTTT 360
TATCCGGAAG ATGAAATTCT TCACTTCTAC AAATGGACGT CGCCCCCAGG GGTGATCCGG 420
ATCCTGTCTA TGCTCATTAT TGTGATGTGC ATCGCCATAT TTGCCTGTGT GGCTTCCACA 480
CTTGCTTGGG ACAGAGGCTA TGGGACAGGG CTCTTTGGAG GAAGCCTAAA CTACCCTTAT 540
AGTGGCTTTG GCTACGGAGG TGGCTATGGA GGCGGCTATG GAGGCTATGG CTATGGCTAT 600
GGCGGATATA CAGACCCAAG AGCAGCCAAA GGCTTCCTGT TGGCCATGGC AGCCTTCTGC 660
TTCATCGCTT CCTTAGTAAT ATTTGTGACC AGTGTTATAA GATCTGCAAT GTCCAGGACA 720
AGAAGATATT ACTTGATCGT GATCATAGTC AGCGCTATCC TGGGCATCAT GGTGTTTATT 780
GCCACGATCG TGTACATAAT GGGAGTGAAC CCGACGGCCC AGGCTTCTGG ATCTATGTAC 840
GGCTCACAGA TATATATGAT CTGCAACCAG TTTTATACTC CTGGAGGTAC TGGTCTCTAC 900
GTGGATCAAT ATTTGTATCA CTACTGTGTG GTTGATCCCC AGGAGGCTAT AGCCATTGTC 960
CTGGGGTTCA TGATTATCGT GGCTTTTGCT TTAATCATCT TTTTTGCTGT GAAAACCCGA 1020
AGAAAGATGG ATCGGTATGA TAAGTCCAAT ATTTTGTGGG ATAAGGAACA CATTTATGAT 1080
GAACAGCCCC CCAATGTTGA AGAGTGGGTT AAAAATGTGT CTGCAGGCAC ACAGGACATG 1140
CCTCCACCCC CATCTGACTA TGCGGAAAGA GTTGACAGTC CAATGGCCTA CTCCTCCAAT 1200
GGCAAAGTGA ATGGCAAGCG ATCATACCCA GAGTCTTTCT ATAAGTCAAC ACCTCTGGTG 1260
CCTGAAGTGG CCCAGGAGAT TCCTCTGACC TTGAGTGTGG ATGACTTCAG GCAGCCTCGG 1320
TACAGCAGCA ATGGTAACCT AGAGACACCT TCTAAAAGGG CTCCCACGAA GGGGAAAGCA 1380
GGAAAGGGCA AGAGGACGGA CCCTGACCAC TATGAAACAG ACTACACGAC AGGTGGGGAG 1440
TCCTGCGAGG AGCTGGAGGA GGACTGGGTC AGGGAATATC CACCTATCAC TTCAGATCAA 1500
CAAAGACAAC TCTACAAGAG AAATTTTGAT GCAGGTCTGC AGGAGTATAA GAGCTTACAG 1560
GCAGAACTAG ACGACGTCAA TAAAGAGCTC TCTCGTCTAG ATAAAGAGCT GGATGACTAC 1620
AGAGAGGAGA GTGAAGAGTA CATGGCTGCT GCTGATGAAT ATAATAGACT AAAGCAAGTT 1680

AAGGGATCTG CAGATTATAA AAGTAAGAGG AATTACTGCA AGCAGTTGAA GAGCAAATTA 1740

TCGCACATCA AGAGGATGGT GGGAGACTAT GACAGACGGA AAGCTTAGAG AGATGCCAGT 1800

TGCGGGAGAA GGGAGAGGTG CATCTGCCTG CACGATGTCT CTGCAATTCT CTCCAGAGGC 1860

AAACTGACTT TGGACTCTAA TCTGGGAAGT TAAAACTTTG TGATCATTAC AAAGTTTCCA 1920

TGGCTTTAAT TCCATCAGTT TCCTATCTCC AGTATTGAAG CATTTTATAA ATGGCTTTTG 1980

ATAATTGACT GGGCTGAACA CTCCAATTAA GGATTTTACA GTTTCAACAT TGATTCTTGT 2040

ATTAAGAATT AAAATGTTGC TTGAGGTTTT AAATGTCAAG AAAGGTCCTG GTGTGAGCTG 2100

TGATGTGTGT GAGCTGTGAT GTGAAGGTTC ACACGCCAGG CAGCGTGTTC CTCCAGGTAG 2160

ACCGTCTAAT CAATCTTTGC AGCAGCCCTC AGGTGACTGT TATTTAGAAT CAGGTTGTTT 2220

TTGGTTTTCC AGACAGGGTT TCTCTGTGTA GCCCTGGCTG ACCTAGAACT TACGCTGTAG 2280

ACCAGGCTGG CCTTGAACTC ACACAGCTCC TCTGAGTGCT GGTGCAGGAG TTAACGTCGT 2340

GGACCGGTAT CATCACTTTT CCTGCGGTGA CTTCTCCAAA CTGAAACTGC TAAGGCAGTT 2400

TTGGCTAAGT CTGTTTTATG ACTGCAAATG ACAGCATTCC TGCCTTTGTA TTTCAGGGCA 2460

AATACGATAC ATTATATCGG CCATGTTCCC CACCACTGTT TTTCTTATAT TGACTTTTAA 2520

CAAATGAATA GGATTATTTT TGGCTTTACA TTTTTTCCTA ACACTTAAGA TCATATAAAA 2580

TTAACAAATA TGTGAAATTT AAGAATTGTA AATATATATT TACGTTTGAA AGATGATTTT 2640

AAATCCAGGG TTAAAGTGCT TTTTATCTTG TATAGTTTAC ATGCTTTTTT TTTTTTTTGA 2700

TAACCCACTA GACCTTTCCA TTGTATCAGA GTATCCAATT ACATTTACAA TTATGACTTG 2760

AATTGTATTT CACAGGAATG CTCAAGTTTT GTACATATTT TATAAGGTAT TAAACCTGAT 2820

GTTCTCTTTC TAAAAAAAA 2839

SEQ ID NO: 7

SEQUENCE LENGTH: 33

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Other nucleic acid Synthesis DNA

SEQUENCE DESCRIPTION

TATGAGACAG ACTACACAAC TGGCGGCGAG TCC

SEQUENCE ID NO: 8

SEQUENCE LENGTH: 30

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Other nucleic acid Synthesis DNA

SEQUENCE DESCRIPTION

ATCATAGTCT CCAACCATCT TCTTGATGTG

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Eisai Co. Ltd.
(B) STREET: 6-10, Koishikawa, 6-chome, Bunkyo-ku
(C) CITY: Tokyo
(E) COUNTRY: Japan
(F) POSTAL CODE (ZIP): none

(ii) TITLE OF INVENTION: Human adhesion molecule occludin

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: EP 97 905 440.0

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: JP 49 880/96
(B) FILING DATE: 07-MAR-1996

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: JP 331 944/96
(B) FILING DATE: 12-DEC-1996

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 522 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(B) STRAIN: human intestinal epithelial cell

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
Met Ser Ser Arg Pro Leu Glu Ser Pro Pro Pro Tyr Arg Pro Asp Glu
1               5                  10                  15

Phe Lys Pro Asn His Tyr Ala Pro Ser Asn Asp Ile Tyr Gly Gly Glu
            20                  25                  30

Met His Val Arg Pro Met Leu Ser Gln Pro Ala Tyr Ser Phe Tyr Pro
        35                  40                  45

Glu Asp Glu Ile Leu His Phe Tyr Lys Trp Thr Ser Pro Pro Gly Val
    50                  55                  60

Ile Arg Ile Leu Ser Met Leu Ile Ile Val Met Cys Ile Ala Ile Phe
65                  70                  75                  80
```

```
Ala Cys Val Ala Ser Thr Leu Ala Trp Asp Arg Gly Tyr Gly Thr Ser
                85                  90                  95

Leu Leu Gly Gly Ser Val Gly Tyr Pro Tyr Gly Gly Ser Gly Phe Gly
            100                 105                 110

Ser Tyr Gly Ser Gly Tyr Gly Tyr Gly Tyr Gly Tyr Gly Tyr Gly Tyr
        115                 120                 125

Gly Gly Tyr Thr Asp Pro Arg Ala Ala Lys Gly Phe Met Leu Ala Met
    130                 135                 140

Ala Ala Phe Cys Phe Ile Ala Ala Leu Val Ile Phe Val Thr Ser Val
145                 150                 155                 160

Ile Arg Ser Glu Met Ser Arg Thr Arg Arg Tyr Tyr Leu Ser Val Ile
                165                 170                 175

Ile Val Ser Ala Ile Leu Gly Ile Met Val Phe Ile Ala Thr Ile Val
            180                 185                 190

Tyr Ile Met Gly Val Asn Pro Thr Ala Gln Ser Ser Gly Ser Leu Tyr
        195                 200                 205

Gly Ser Gln Ile Tyr Ala Leu Cys Asn Gln Phe Tyr Thr Pro Ala Ala
    210                 215                 220

Thr Gly Leu Tyr Val Asp Gln Tyr Leu Tyr His Tyr Cys Val Val Asp
225                 230                 235                 240

Pro Gln Glu Ala Ile Ala Ile Val Leu Gly Phe Met Ile Ile Val Ala
                245                 250                 255

Phe Ala Leu Ile Ile Phe Phe Ala Val Lys Thr Arg Arg Lys Met Asp
            260                 265                 270

Arg Tyr Asp Lys Ser Asn Ile Leu Trp Asp Lys Glu His Ile Tyr Asp
        275                 280                 285

Glu Gln Pro Pro Asn Val Glu Glu Trp Val Lys Asn Val Ser Ala Gly
    290                 295                 300

Thr Gln Asp Val Pro Ser Pro Pro Ser Asp Tyr Val Glu Arg Val Asp
305                 310                 315                 320

Ser Pro Met Ala Tyr Ser Ser Asn Gly Lys Val Asn Asp Lys Arg Phe
                325                 330                 335

Tyr Pro Glu Ser Ser Tyr Lys Ser Thr Pro Val Pro Glu Val Val Gln
            340                 345                 350

Glu Leu Pro Leu Thr Ser Pro Val Asp Asp Phe Arg Gln Pro Arg Tyr
        355                 360                 365

Ser Ser Gly Gly Asn Phe Glu Thr Pro Ser Lys Arg Ala Pro Ala Lys
    370                 375                 380

Gly Arg Ala Gly Arg Ser Lys Arg Thr Glu Gln Asp His Tyr Glu Thr
385                 390                 395                 400

Asp Tyr Thr Thr Gly Gly Glu Ser Cys Asp Glu Leu Glu Glu Asp Trp
                405                 410                 415

Ile Arg Glu Tyr Pro Pro Ile Thr Ser Asp Gln Gln Arg Gln Leu Tyr
            420                 425                 430
```

```
Lys Arg Asn Phe Asp Thr Gly Leu Gln Glu Tyr Lys Ser Leu Gln Ser
        435                 440                 445

Glu Leu Asp Glu Ile Asn Lys Glu Leu Ser Arg Leu Asp Lys Glu Leu
    450                 455                 460

Asp Asp Tyr Arg Glu Glu Ser Glu Glu Tyr Met Ala Ala Ala Asp Glu
465                 470                 475                 480

Tyr Asn Arg Leu Lys Gln Val Lys Gly Ser Ala Asp Tyr Lys Ser Lys
                485                 490                 495

Lys Asn His Cys Lys Gln Leu Lys Ser Lys Leu Ser His Ile Lys Lys
            500                 505                 510

Met Val Gly Asp Tyr Asp Arg Gln Lys Thr
        515                 520
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 521 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: canine
(B) STRAIN: canine kidney cell MDCK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Ser Ser Arg Pro Phe Glu Ser Pro Pro Pro Tyr Arg Pro Asp Glu
1               5                   10                  15

Phe Lys Pro Asn His Tyr Ala Pro Ser Asn Asp Val Tyr Gly Gly Asp
            20                  25                  30

Met His Val Arg Pro Met Leu Ser Gln Pro Ala Tyr Ser Phe Tyr Pro
        35                  40                  45

Glu Asp Glu Ile Leu His Phe Tyr Lys Trp Thr Ser Pro Pro Gly Val
    50                  55                  60

Ile Arg Ile Leu Ser Met Leu Val Ile Val Met Cys Ile Ala Ile Phe
65                  70                  75                  80

Gly Cys Val Ala Ser Thr Leu Ala Trp Asp Arg Gly Tyr Gly Thr Gly
                85                  90                  95

Leu Met Gly Gly Ser Ile Gly Tyr Pro Tyr Gly Ser Gly Phe Gly Ser
            100                 105                 110

Tyr Gly Thr Gly Tyr Gly Tyr Gly Phe Gly Tyr Gly Tyr Gly Tyr Gly
        115                 120                 125

Gly Tyr Thr Asp Pro Arg Ala Ala Lys Gly Phe Leu Leu Ala Met Val
    130                 135                 140

Ala Phe Cys Phe Ile Ala Ala Leu Val Ile Phe Val Thr Ser Val Ile
145                 150                 155                 160
```

```
Arg Ser Asp Ile Ser Arg Thr Arg Arg Tyr Tyr Leu Thr Val Ile Ile
                    165                 170                 175

Leu Ser Ala Phe Leu Gly Val Met Met Phe Ile Ala Thr Ile Val Tyr
                180                 185                 190

Ile Met Gly Val Asn Pro Thr Ala Gln Ala Ser Gly Ser Leu Tyr Ser
            195                 200                 205

Ser Gln Ile Tyr Ala Met Cys Asn Gln Phe Tyr Ala Ser Thr Ala Thr
    210                 215                 220

Gly Leu Tyr Met Asp Gln Tyr Leu Tyr His Tyr Cys Val Val Asp Pro
225                 230                 235                 240

Gln Glu Ala Ile Ala Ile Val Leu Gly Phe Met Val Ile Val Ala Phe
                245                 250                 255

Ala Leu Ile Ile Phe Phe Ala Val Lys Thr Arg Arg Lys Met Asp Arg
                260                 265                 270

Tyr Asp Lys Ser Asn Ile Leu Trp Asp Lys Glu His Ile Tyr Asp Glu
            275                 280                 285

Gln Pro Pro Asn Val Glu Glu Trp Val Lys Asn Val Ser Ala Gly Thr
    290                 295                 300

Gln Asp Met Pro Pro Pro Pro Ser Asp Tyr Val Glu Arg Val Asp Ser
305                 310                 315                 320

Pro Met Ala Tyr Ser Ser Asn Gly Lys Val Asn Asp Lys Arg Leu Tyr
                325                 330                 335

Pro Glu Ser Ser Tyr Lys Ser Thr Pro Val Pro Glu Val Val Gln Glu
                340                 345                 350

Leu Pro Ala Thr Ser Pro Ala Asp Asp Phe Arg Gln Pro Arg Tyr Ser
            355                 360                 365

Ser Ser Gly His Leu Glu Pro Pro Ser Lys Arg Ala Pro Ser Lys Gly
    370                 375                 380

Arg Thr Gly Arg Pro Lys Arg Leu Glu Gln Asp His Tyr Glu Thr Asp
385                 390                 395                 400

Tyr Thr Thr Gly Gly Glu Ser Cys Asp Glu Leu Glu Glu Asp Trp Ile
                405                 410                 415

Arg Glu Tyr Pro Pro Ile Thr Ser Asp Gln Gln Arg Gln Leu Tyr Lys
                420                 425                 430

Arg Asn Phe Asp Thr Gly Leu Gln Glu Tyr Lys Ser Leu Gln Ala Glu
            435                 440                 445

Leu Asp Glu Ile Asn Lys Glu Leu Ser Arg Leu Asp Lys Glu Leu Asp
    450                 455                 460

Asp Tyr Arg Glu Glu Ser Glu Glu Tyr Met Ala Ala Ala Asp Glu Tyr
465                 470                 475                 480

Asn Arg Leu Lys Gln Val Lys Gly Ser Pro Asp Tyr Lys Asn Lys Arg
                485                 490                 495

Asn Tyr Cys Lys Gln Leu Lys Ser Lys Leu Ser His Ile Lys Lys Met
                500                 505                 510
```

Val Gly Asp Tyr Asp Arg Gln Lys Thr
515 520

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 521 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mouse
(B) STRAIN: mouse lung cell

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Ser Val Arg Pro Phe Glu Ser Pro Pro Pro Tyr Arg Pro Asp Glu
1 5 10 15

Phe Lys Pro Asn His Tyr Ala Pro Ser Asn Asp Met Tyr Gly Gly Glu
20 25 30

Met His Val Arg Pro Met Leu Ser Gln Pro Ala Tyr Ser Phe Tyr Pro
35 40 45

Glu Asp Glu Ile Leu His Phe Tyr Lys Trp Thr Ser Pro Pro Gly Val
50 55 60

Ile Arg Ile Leu Ser Met Leu Ile Ile Val Met Cys Ile Ala Ile Phe
65 70 75 80

Ala Cys Val Ala Ser Thr Leu Ala Trp Asp Arg Gly Tyr Gly Thr Gly
85 90 95

Leu Phe Gly Gly Ser Leu Asn Tyr Pro Tyr Ser Gly Phe Gly Tyr Gly
100 105 110

Gly Gly Tyr Gly Gly Gly Tyr Gly Gly Tyr Gly Tyr Gly Tyr Gly Gly
115 120 125

Tyr Thr Asp Pro Arg Ala Ala Lys Gly Phe Leu Leu Ala Met Ala Ala
130 135 140

Phe Cys Phe Ile Ala Ser Leu Val Ile Phe Val Thr Ser Val Ile Arg
145 150 155 160

Ser Gly Met Ser Arg Thr Arg Arg Tyr Tyr Leu Ile Val Ile Ile Val
165 170 175

Ser Ala Ile Leu Gly Ile Met Val Phe Ile Ala Thr Ile Val Tyr Ile
180 185 190

Met Gly Val Asn Pro Thr Ala Gln Ala Ser Gly Ser Met Tyr Gly Ser
195 200 205

Gln Ile Tyr Met Ile Cys Asn Gln Phe Tyr Thr Pro Gly Gly Thr Gly
210 215 220

Leu Tyr Val Asp Gln Tyr Leu Tyr His Tyr Cys Val Val Asp Pro Gln
225 230 235 240

OFFICIAL COMMUNICATION

SEED
Intellectual Property Law Group PLLC

701 Fifth Avenue, Suite 6300
Seattle WA 98104-7092 USA
Facsimile: (206) 682-6031
Telephone: (206) 622-4900
www.seedlaw.com

Qing Lin, Ph.D.
beckyl@seedlaw.com

May 1, 2002

Facsimile Transmission 2704 **pages (including this page)**

TO: U.S. Patent and Trademark Office – SPE Christopher Low
Fax No.: 1-703-746-5040

RE: Application serial no.: 09/450,073
Art unit: 1653
Filing date of application: November 29, 1999
Examiner assigned to application: Avis M. Davenport
Title of the invention: COMPOUNDS AND METHODS FOR CANCER THERAPY
Attorney docket number: 100086.405C2

☒ **Urgent** ☐ For Review ☒ Please Confirm Receipt ☐ Please Reply ASAP

Comments:

PLEASE DELIVER IMMEDIATELY TO

EXAMINER: CHRISTOPHER LOW

As discussed with Judi Breaks of my office, enclosed are the reference for the IDS filed in this matter on 8-2-00. Please let us know if you need anything in addition.

If you do not receive all pages, please call Judi Breaks at (206) 622-4900 or fax our office.

Transmission Information: Date _____ Time _____ By _____
D:\NrPortbl\iManage\JUDIB\215246_1.DOC

Glu Ala Ile Ala Ile Val Leu Gly Phe Met Ile Ile Val Ala Phe Ala
245 250 255

Leu Ile Ile Phe Phe Ala Val Lys Thr Arg Arg Lys Met Asp Arg Tyr
260 265 270

Asp Lys Ser Asn Ile Leu Trp Asp Lys Glu His Ile Tyr Asp Glu Gln
275 280 285

Pro Pro Asn Val Glu Glu Trp Val Lys Asn Val Ser Ala Gly Thr Gln
290 295 300

Asp Met Pro Pro Pro Pro Ser Asp Tyr Ala Glu Arg Val Asp Ser Pro
305 310 315 320

Met Ala Tyr Ser Ser Asn Gly Lys Val Asn Gly Lys Arg Ser Tyr Pro
325 330 335

Glu Ser Phe Tyr Lys Ser Thr Pro Leu Val Pro Glu Val Ala Gln Glu
340 345 350

Ile Pro Leu Thr Leu Ser Val Asp Asp Phe Arg Gln Pro Arg Tyr Ser
355 360 365

Ser Asn Gly Asn Leu Glu Thr Pro Ser Lys Arg Ala Pro Thr Lys Gly
370 375 380

Lys Ala Gly Lys Gly Lys Arg Thr Asp Pro Asp His Tyr Glu Thr Asp
385 390 395 400

Tyr Thr Thr Gly Gly Glu Ser Cys Glu Glu Leu Glu Glu Asp Trp Val
405 410 415

Arg Glu Tyr Pro Pro Ile Thr Ser Asp Gln Gln Arg Gln Leu Tyr Lys
420 425 430

Arg Asn Phe Asp Ala Gly Leu Gln Glu Tyr Lys Ser Leu Gln Ala Glu
435 440 445

Leu Asp Asp Val Asn Lys Glu Leu Ser Arg Leu Asp Lys Glu Leu Asp
450 455 460

Asp Tyr Arg Glu Glu Ser Glu Glu Tyr Met Ala Ala Ala Asp Glu Tyr
465 470 475 480

Asn Arg Leu Lys Gln Val Lys Gly Ser Ala Asp Tyr Lys Ser Lys Arg
485 490 495

Asn Tyr Cys Lys Gln Leu Lys Ser Lys Leu Ser His Ile Lys Arg Met
500 505 510

Val Gly Asp Tyr Asp Arg Arg Lys Pro
515 520

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2379 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens
(B) STRAIN: human intestinal epithelial cell strain T84

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:168..1733

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CTCCCGCGTC CACCTCTCCC TCCCTGCTTC CTCTGGCGGA GGCGGCAGGA ACCGAGAGCC 60
AGGTCCAGAG CGCCGAGGAG CCGGTCTAGG ACGCAGCAGA TTGGTTTATC TTGGAAGCTA 120
AAGGGCATTG CTCATCCTGA AGATCAGCTG ACCATTGACA ATCAGCCATG TCATCCAGGC 180
CTCTTGAAAG TCCACCTCCT TACAGGCCTG ATGAATTCAA ACCGAATCAT TATGCACCJA 240
GCAATGACAT ATATGGTGGA GAGATGCATG TTCGACCAAT GCTCTCTCAG CCAGCCTACT 300
CTTTTTACCC AGAAGATGAA ATTCTTCACT TCTACAAATG GACCTCTCCT CCAGGAGTGA 360
TTCGGATCCT GTCTATGCTC ATTATTGTGA TGTGCATTGC CATCTTTGCC TGTGTGGCCT 420
CCACGCTTGC CTGGGACAGA GGCTATGGAA CTTCCCTTTT AGGAGGTAGT GTAGGCTACC 480
CTTATGGAGG AAGTGGCTTT GGTAGCTACG GAAGTGGCTA TGGCTATGGC TATGGTTATG 540
GCTATGGCTA CGGAGGCTAT ACAGACCCAA GAGCAGCAAA GGGCTTCATG TTGGCCATGG 600
CTGCCTTTTG TTTCATTGCC GCGTTGGTGA TCTTTGTTAC CAGTGTTATA AGATCTGAAA 660
TGTCCAGAAC AAGAAGATAC TACTTAAGTG TGATAATAGT GAGTGCTATC CTGGGCATCA 720
TGGTGTTTAT TGCCACAATT GTCTATATAA TGGGAGTGAA CCCAACTGCT CAGTCTTCTG 780
GATCTCTATA TGGTTCACAA ATATATGCCC TCTGCAACCA ATTTTATACA CCTGCAGCTA 840
CTGGACTCTA CGTGGATCAG TATTTGTATC ACTACTGTGT TGTGGATCCC CAGGAGGCCA 900
TTGCCATTGT ACTGGGGTTC ATGATTATTG TGGCTTTTGC TTTAATAATT TTCTTTGCTG 960
TGAAAACTCG AAGAAAGATG GACAGGTATG ACAAGTCCAA TATTTTGTGG GACAAGGAAC 1020
ACATTTATGA TGAGCAGCCC CCCAATGTCG AGGAGTGGGT TAAAAATGTG TCTGCAGGCA 1080
CACAGGACGT GCCTTCACCC CCATCTGACT ATGTGGAAAG AGTTGACAGT CCCATGGCAT 1140
ACTCTTCCAA TGGCAAAGTG AATGACAAGC GGTTTTATCC AGAGTCTTCC TATAAATCCA 1200
CGCCGGTTCC TGAAGTGGTT CAGGAGCTTC CATTAACTTC GCCTGTGGAT GACTTCAGGC 1260
AGCCTCGTTA CAGCAGCGGT GGTAACTTTG AGACACCTTC AAAAAGAGCA CCTGCAAAGG 1320
GAAGAGCAGG AAGGTCAAAG AGAACAGAGC AAGATCACTA TGAGACAGAC TACACAACTG 1380
GCGGCGAGTC CTGTGATGAG CTGGAGGAGG ACTGGATCAG GGAATATCCA CCTATCACTT 1440
CAGATCAACA AAGACAACTG TACAAGAGGA ATTTTGACAC TGGCCTACAG GAATACAAGA 1500
GCTTACAATC AGAACTTGAT GAGATCAATA AAGAACTCTC CCGTTTGGAT AAAGAATTGG 1560
ATGACTATAG AGAAGAAAGT GAAGAGTACA TGGCTGCTGC TGATGAATAC AATAGACTGA 1620
AGCAAGTGAA GGGATCTGCA GATTACAAAA GTAAGAAGAA TCATTGCAAG CAGTTAAAGA 1680

GCAAATTGTC ACACATCAAG AAGATGGTTG GAGACTATGA TAGACAGAAA ACATAGAAGG 1740

CTGATGCCAA GTTGTTTGAG AAATTAAGTA TCTGACATCT CTGCAATCTT CTCAGAAGGC 1800

AAATGACTTT GGACCATAAC CCCGGAAGCC AAACCTCTGT GAGCATCACA AAGTTTTGGT 1860

TGCTTTAACA TCATCAGTAT TGAAGCATTT TATAAATCGC TTTTGATAAT CAACTGGGCT 1920

GAACACTCCA ATTAAGGATT TTATGCTTTA AACATTGGTT CTTGTATTAA GAATGAAATA 1980

CTGTTTGAGG TTTTTAAGCC TTAAAGGAAG GTTCTGGTGT GAACTAAACT TTCACACCCC 2040

AGACGATGTC TTCATACCTA CATGTATTTG TTTGCATAGG TGATCTCATT TAATCCTCTC 2100

AACCACCTTT CAGATAACTG TTATTTATAA TCACTTTTTT CCACATAAGG AAACTGGGTT 2160

CCTGCAATGA AGTCTCTGAA GTGAAACTGC TTGTTTCCTA GCACACACTT TTGGTTAAGT 2220

CTGTTTTATG ACTTCATTAA TAATAAATTC CCTGGCCTTT CATATTTTAG CTACTATATA 2280

TGTGATGATC TACCAGCCTC CCTATTTTTT TTCTGTTATA TAAATGGTTA AAAGAGGTTT 2340

TTCTTAAATA ATAAAGATCA TGTAAAAGTA AAAAAAAAA 2379

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1961 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: canine
(B) STRAIN: canine kidney cell strain MDCK

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:72..1624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CAGGTTGGCT TATTTTGGGG AGCTCTGGGA TCCTGCTCGT CCTGAAGATC GGGTGATCAT 60

TGACATCAGC CATGTCATCG AGGCCTTTTG AGAGTCCACC TCCGTATAGA CCTGATGAAT 120

TCAAACCCAA TCATTATGCA CCGAGCAATG ATGTGTACGG TGGGGACATG CACGTCCGAC 180

CCATGCTCTC TCAGCCGGCG TATTCTTTCT ACCCAGAAGA TGAAATTCTT CACTTCTACA 240

AATGGACCTC TCCTCCAGGA GTAATTCGGA TTCTGTCCAT GCTTGTCATT GTGATGTGCA 300

TCGCCATATT TGGCTGTGTC GCGTCCACGC TCGCCTGGGA TAGAGGCTAT GGAACTGGCT 360

TAATGGGTGG TAGCATAGGC TACCCTTACG GAAGTGGCTT CGGGAGCTAC GGGACTGGCT 420

ACGGCTACGG GTTTGGCTAC GGCTACGGCT ACGGCGGCTA CACGGATCCC AGAGCAGCAA 480

AGGGCTTCCT CCTGGCCATG GTGGCCTTTT GTTTTATCGC TGCATTGGTG ATATTTGTTA 540

CCAGCGTTAT AAGGTCTGAC ATATCCAGAA CCAGAAGGTA CTACTTGACT GTAATAATAC 600

TGAGTGCCTT CCTGGGCGTC ATGATGTTCA TTGCTACAAT TGTCTATATA ATGGGAGTCA 660

ATCCAACTGC CCAGGCTTCT GGGTCTTTAT ACAGTTCACA GATATATGCC ATGTGCAACC 720

AGTTCTATGC ATCTACAGCT ACCGGACTCT ACATGGATCA GTATTTGTAT CACTACTGTG 780

TGGTGGATCC CCAAGAGGCA ATTGCCATTG TCCTGGGATT CATGGTGATT GTGGCTTTTG 840

CTTTAATAAT TTTCTTTGCT GTGAAAACTC GAAGAAAGAT GGACCGGTAT GACAAGTCGA 900

ATATATTGTG GGACAAGGAA CATATTTATG ATGAACAACC CCCCAATGTT GAAGAGTGGG 960

TTAAAAACGT TTCTGCAGGC ACACAAGACA TGCCTCCTCC CCCTTCTGAC TATGTGGAGA 1020

GAGTGGACAG TCCCATGGCG TACTCTTCCA ATGGTAAAGT GAATGACAAG CGGTTGTATN 1080

CAGAGTCTTC CTATAAATCA ACACCGGTCC CCGAAGTGGT GCAGGAGCTG CCCGCCACCT 1140

CCCCTGCGGA TGACTTCAGG CAGCCTCGCT ACAGCAGCAG CGGGCACTTG GAGCCACCTT 1200

CGAAGAGGGC CCCCTCGAAA GGAAGAACGG GAAGGCCCAA GAGGCTGGAG CAGGACCACT 1260

ATGAGACAGA CTACACGACG GGCGGCGAGT CGTGTGACGA GCTGGAGGAG GACTGGATCA 1320

GGGAATATCC ACCTATCACT TCAGATCAAC AAAGACAACT CTACAAGAGA AATTTTGACA 1380

CTGGCCTGCA GGAATACAAG AGCTTACAAG CAGAACTTGA TGAGATCAAT AAAGAACTCT 1440

CTCGCCTGGA TAAAGAATTG GATGACTATA GAGAAGAAAG TGAAGAGTAC ATGGCTGCTG 1500

CTGATGAGTA CAATAGACTG AAGCAAGTTA AGGGATCTCC AGATTACAAA AATAAGAGGA 1560

ATTATTGCAA GCAGTTGAAG AGCAAATTGT CCCACATCAA GAAGATGGTT GGAGACTATG 1620

ATAGACAGAA AACATAGAAG GCAGATGCCA CACAGTTTGA GAGATTGTGA AGTATTTGAC 1680

ATATCTGCAA CGTTGTCAGA AGGCAGAATG ACTTTGGATT TCGAACCCAG GAGGCCAGAT 1740

CTTTGTGATC ATTACAAAGT TTTGGTAGCT TTAATATCAT CAGTATTGAA GCATTTTACA 1800

CATAGCTTTT GATAATCAAC TGGGCTGAAC ACTCCCGATT AAGGATTCTG TGCTTTAGAC 1860

TTTGGCTGTT GTGCTAAAGG ACTGAGTATA GGTGGAGGTT TTCAGACCTT GGAAGAAGGT 1920

CCCACGGTGA ACTTGTGCTG TGAACTTGCA CACTTGGGGC A 1961

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2839 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mouse
(B) STRAIN: mouse lung cell

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION:223..1785

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
GGAGTTTCAG GTGAATGGGT CACCGAGGGA GGAGGCTGGC CACGCCACAC CTCGTCGCTA    60
GTGCCCACCT CCCGGCCCCT CTTTCCTTAG GCGACAGCGG TGGAGTTGCG GGAGAGCGGT   120
CCAGCGCACG GAGCAACCGG CTAGGGGCTC GGCAGGTTCG CTTATCTTGG GAGCCTGGAC   180
ATTTTGCTCA TCATAAAGAT TAGGTGACCA GTGACATCAG CCATGTCCGT GAGGCCTTTT   240
GAAAGTCCAC CTCCTTACAG ACCTGATGAA TTCAAACCCA ATCATTATGC ACCAAGCAAT   300
GACATGTATG GCGGAGAGAT GCATGTCCGG CCGATGCTCT CTCAGCCAGC GTACTCTTTT   360
TATCCGGAAG ATGAAATTCT TCACTTCTAC AAATGGACGT CGCCCCCAGG GGTGATCCGG   420
ATCCTGTCTA TGCTCATTAT TGTGATGTGC ATCGCCATAT TTGCCTGTGT GGCTTCCACA   480
CTTGCTTGGG ACAGAGGCTA TGGGACAGGG CTCTTTGGAG GAAGCCTAAA CTACCCTTAT   540
AGTGGCTTTG GCTACGGAGG TGGCTATGGA GGCGGCTATG GAGGCTATGG CTATGGCTAT   600
GGCGGATATA CAGACCCAAG AGCAGCCAAA GGCTTCCTGT TGGCCATGGC AGCCTTCTGC   660
TTCATCGCTT CCTTAGTAAT ATTTGTGACC AGTGTTATAA GATCTGGAAT GTCCAGGACA   720
AGAAGATATT ACTTGATCGT GATCATAGTC AGCGCTATCC TGGGCATCAT GGTGTTTATT   780
GCCACGATCG TGTACATAAT GGGAGTGAAC CCGACGGCCC AGGCTTCTGG ATCTATGTAC   840
GGCTCACAGA TATATATGAT CTGCAACCAG TTTTATACTC CTGGAGGTAC TGGTCTCTAC   900
GTGGATCAAT ATTTGTATCA CTACTGTGTG GTTGATCCCC AGGAGGCTAT AGCCATTGTC   960
CTGGGGTTCA TGATTATCGT GGCTTTTGCT TTAATCATCT TTTTTGCTGT GAAAACCCGA  1020
AGAAAGATGG ATCGGTATGA TAAGTCCAAT ATTTTGTGGG ATAAGGAACA CATTTATGAT  1080
GAACAGCCCC CCAATGTTGA AGAGTGGGTT AAAAATGTGT CTGCAGGCAC ACAGGACATG  1140
CCTCCACCCC CATCTGACTA TGCGGAAAGA GTTGACAGTC CAATGGCCTA CTCCTCCAAT  1200
GGCAAAGTGA ATGGCAAGCG ATCATACCCA GAGTCTTTCT ATAAGTCAAC ACCTCTGGTG  1260
CCTGAAGTGG CCCAGGAGAT TCCTCTGACC TTGAGTGTGG ATGACTTCAG GCAGCCTCGG  1320
TACAGCAGCA ATGGTAACCT AGAGACACCT TCTAAAAGGG CTCCCACGAA GGGGAAAGCA  1380
GGAAAGGGCA AGAGGACGGA CCCTGACCAC TATGAAACAG ACTACACGAC AGGTGGGGAG  1440
TCCTGCGAGG AGCTGGAGGA GGACTGGGTC AGGGAATATC CACCTATCAC TTCAGATCAA  1500
CAAAGACAAC TCTACAAGAG AAATTTTGAT GCAGGTCTGC AGGAGTATAA GAGCTTACAG  1560
GCAGAACTAG ACGACGTCAA TAAAGAGCTC TCTCGTCTAG ATAAAGAGCT GGATGACTAC  1620
AGAGAGGAGA GTGAAGAGTA CATGGCTGCT GCTGATGAAT ATAATAGACT AAAGCAAGTT  1680
AAGGGATCTG CAGATTATAA AAGTAAGAGG AATTACTGCA AGCAGTTGAA GAGCAAATTA  1740
TCGCACATCA AGAGGATGGT GGGAGACTAT GACAGACGGA AACCTTAGAG AGATGCCAGT  1800
TGCGGGAGAA GGGAGAGGTG CATCTGCCTG CACGATGTCT CTGCAATTCT CTCCAGAGGC  1860
AAACTGACTT TGGACTCTAA TCTGGGAAGT TAAAACTTTG TGATCATTAC AAAGTTTCCA  1920
TGGCTTTAAT TCCATCAGTT TCCTATCTCC AGTATTGAAG CATTTTATAA ATGGCTTTTG  1980
```

ATAATTGACT GGGCTGAACA CTCCAATTAA GGATTTTACA GTTTCAACAT TGATTCTTGT 2040

ATTAAGAATT AAAATGTTGC TTGAGGTTTT AAATGTCAAG AAAGGTCCTG GTGTGAGCTG 2100

TGATGTGTGT GAGCTGTGAT GTGAAGGTTC ACACGCCAGG CAGCGTGTTC CTCCAGGTAG 2160

ACCGTCTAAT CAATCTTTGC AGCAGCCCTC AGGTGACTGT TATTTAGAAT CAGGTTGTTT 2220

TTGGTTTTCC AGACAGGGTT TCTCTGTGTA GCCCTGGCTG ACCTAGAACT TACGCTGTAG 2280

ACCAGGCTGG CCTTGAACTC ACACAGCTCC TCTGAGTGCT GGTGCAGGAG TTAACGTCGT 2340

GGACCGGTAT CATCACTTTT CCTGCGGTGA CTTCTCCAAA CTGAAACTGC TAAGGCAGTT 2400

TTGGCTAAGT CTGTTTTATG ACTGCAAATG ACAGCATTCC TGCCTTTGTA TTTCAGGGCA 2460

AATACGATAC ATTATATCGG CCATGTTCCC CACCACTGTT TTTCTTATAT TGACTTTTAA 2520

CAAATGAATA GGATTATTTT TGGCTTTACA TTTTTTCCTA ACACTTAAGA TCATATAAAA 2580

TTAACAAATA TGTGAAATTT AAGAATTGTA AATATATATT TACGTTTGAA AGATGATTTT 2640

AAATCCAGGG TTAAAGTGCT TTTTATCTTG TATAGTTTAC ATGCTTTTTT TTTTTTTTGA 2700

TAACCCACTA GACCTTTCCA TTGTATCAGA GTATCCAATT ACATTTACAA TTATGACTTC 2760

AATTGTATTT CACAGGAATG CTCAAGTTTT GTACATATTT TATAAGGTAT TAAACCTGAT 2820

GTTCTCTTTC TAAAAAAAA 2839

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthesis DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TATGAGACAG ACTACACAAC TGGCGGCGAG TCC 33

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "Synthesis DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

ATCATAGTCT CCAACCATCT TCTTGATGTG 30

Brief Description of the Drawings:

Fig. 1 is photographs substituting for drawings showing the morphology of organisms. Photo. A is a stained immunofluorescent photomicrograph of cultured human intestinal epithelial cell strain T84 with an anti-human occludin rat monoclonal antibody. Photo. B is a stained immunofluorescent photomicrograph of cultured human intestinal epithelial cell strain T84 with a mouse monoclonal antibody against the TJ lining protein ZO-1. The same sites of T84 were photographed.

Claims

1. A DNA for encoding a human occludin protein having an amino acid sequence as defined in SEQ ID NO. 1.

2. The DNA for encoding a human occludin protein according to claim 1 which is a DNA defined in SEQ ID NO. 4.

3. A DNA for encoding a protein having an amino acid sequence in which one or plural amino acids in an amino acid sequence defined in SEQ ID NO. 1 are added, deleted or substituted.

4. A DNA defined in Sequence No. 5 for encoding a canine occludine protein having an amino acid sequence defined in SEQ ID NO. 2.

5. A DNA defined in SEC ID NO. 6 for encoding a mouse occludin protein having an amino acid sequence defined in SEQ ID NO. 3.

6. A vector which comprises the DNA described in any one of claims 1 to 5.

7. A transformant which holds the vector described in claim 6.

8. A human occludin protein which has an amino acid sequence defined in SEQ ID NO. 1.

9. A protein having an amino acid sequence in which one or plural amino acids in an amino acid sequence defined in SEQ ID NO. 1 are added, deleted or substituted.

10. A partial peptide of a human occludin protein which has an amino acid sequence defined in SEQ ID NO. 1.

11. A canine occludin protein which has an amino acid sequence defined in SEQ ID NO. 2.

12. A mouse occludin protein which has an amino acid sequence defined in SEQ ID NO. 3.

13. A method of manufacturing the protein described in any one of claims 8 to 11 which comprises the steps of cultivating the transformant described in Claim 7 and collecting an expressed product.

14. A DNA probe which comprises all or a part of a base sequence defined in SEQ ID NO. 4, 5 or 3.

15. A DNA primer which comprises a part of a base sequence as defined in SEQ ID NO. 4, 5 or 6.

16. A polyclonal antibody or a monoclonal antibody which specifically binds to the protein described in any one of claims 8, 11 and 12.

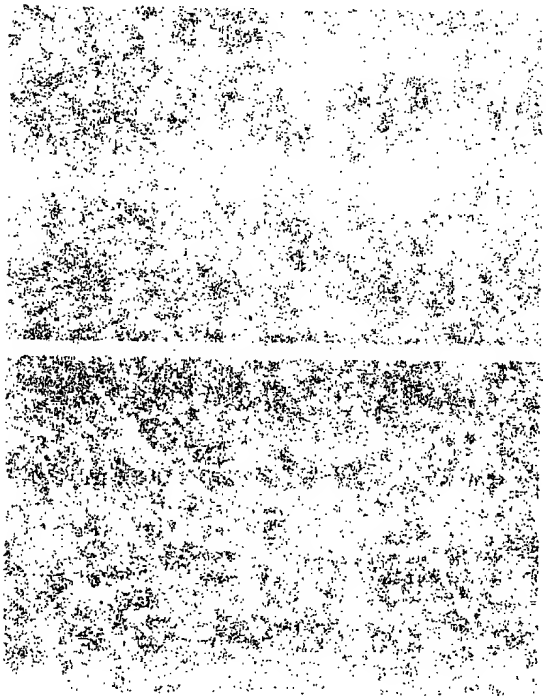
17. An analysis method of the DNA gene described in claim 2, 4 or 5 in a biological specimen, wherein the DNA primer described in claim 15 is used.

18. An analysis method of the DNA gene described in claim 2, 4 or 5 in a biological specimen, wherein the DNA probe described in claim 14 is used.

19. A screening method of a drug affecting the expression of occludin, which comprises the steps of allowing occludin-expressing cells and an analyte to coexist, and then determining an expression quantity of an occludin gene of the cells by the method of claim 17 or 18.

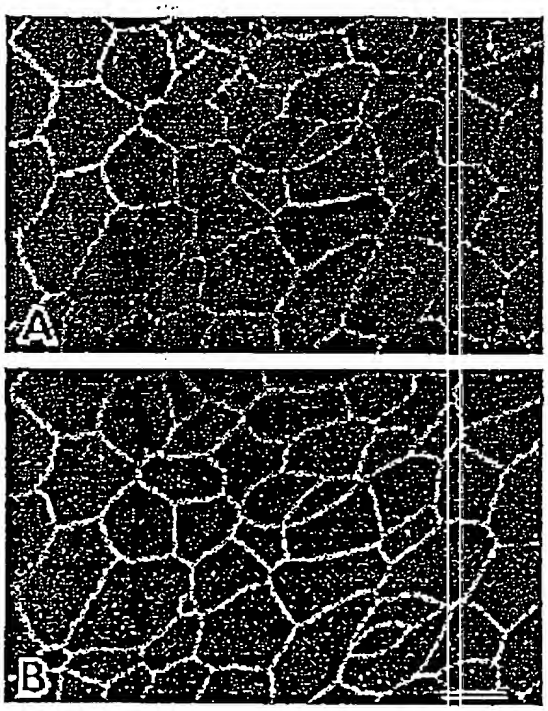
20. A polynucleotide for medical use which comprises at least 6 nucleotides for selectively hybridizing the human occludin DNA described in claim 4.

21. An assay reagent for occludin in a biological specimen, which comprises the antibody described in claim 16.



EP 0 831 148 A1

Fig. 1



EP 0 831 148 A1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP97/00665

A. CLASSIFICATION OF SUBJECT MATTER Int. Cl[6] C12N15/12, C12N15/63, C07K14/435, C12N1/21, C12P21/02, C12Q1/68, C07K16/18, C12P21/08, A61K38/17, G01N33/53 // (C12N1/21, C12R1:19), (C12P21/02, C12R1:19),
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int. Cl[6] C12N15/12, C12N15/63, C07K14/435, C12N1/21, C12P21/02, C12Q1/68, C07K16/18, C12P21/08, A61K38/17, G01N33/53

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPIL
BIOSIS PREVIEWS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X // A	Cell 80 (Jan. 1995) Natalie Roy et al. "The Gene for Neural Apoptosis Inhibitory Protein Is Partially Deleted in Individuals with Spinal Muscular Atrophy" p. 167-178	3, 6, 7, 9, 13-15 // 1, 2, 4, 5, 8, 10-12, 16-21
P,X	J. Cell Biol. 133(1) (Apr. 1996) Yuhko Ando-Akatsuka et al. "Interspecies Diversity of the Occludin Sequence: cDNA Cloning of Human, Mouse, Dog, and Rat-Kangaroo Homologues" p. 43-47	1 - 21
A	J. Cell. Biol. 123(6) (1993) Mikio Furuse et al. "Occludin: A Novel Integral Membrane Protein Localizing at Tight Junctions: p. 1777-1788	1 - 21

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
June 3, 1997 (03. 06. 97)	June 10, 1997 (10. 06. 97)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)